

1473-071

U.S. APPLICATION NO. (If known, see 37 CFR 1.5)

APPLICATION NO. (If known, see 37 CFR  
107088334

INTERNATIONAL FILING DATE

**25 October 2000 (25.10.2000)**

PRIORITY DATE CLAIMED

**26 October 1999 (26.10.1999)**

## EMPHASIS OF SHORT-DURATION TRANSIENT SPEECH FEATURES

APPLICANT(S) FOR DO/EO/US

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. ☐ This is an express request to begin national examination procedures (35 U.S.C. 371(f)). The submission must include items (5), (6), (9) and (21) indicated below.
4. ☐ The US has been elected by the expiration of 19 months from the priority date (Article 31).
5. ☐ A copy of the International Application as filed (35 U.S.C. 371(c)(2))
  - a. ☐ is attached hereto (required only if not communicated by the International Bureau).
  - b. ☐ has been communicated by the International Bureau.
  - c. ☐ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. ☐ An English language translation of the International Application as filed (35 U.S.C. 371(c)(2)).
  - a. ☐ is attached hereto.
  - b. ☐ has been previously submitted under 35 U.S.C. 154(d)(4).
7. ☒ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
  - a. ☐ are attached hereto (required only if not communicated by the International Bureau).
  - b. ☐ have been communicated by the International Bureau.
  - c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
  - d. ☒ have not been made and will not be made.
8. ☐ An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371 (c)(3)).
9. ☐ An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
10. ☒ An English language translation of the annexes of the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

**Items 11 to 20 below concern document(s) or information included:**

11. ☒ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
13. ☒ A FIRST preliminary amendment.
14. ☐ A SECOND or SUBSEQUENT preliminary amendment.
15. ☐ A substitute specification.
16. ☐ A change of power of attorney and/or address letter.
17. ☐ A computer-readable form of the sequence listing in accordance with PCT Rule 13ter.2 and 35 U.S.C. 1.821 - 1.825.
18. ☒ A second copy of the published international application under 35 U.S.C. 154(d)(4).
19. ☐ A second copy of the English language translation of the international application under 35 U.S.C. 154(d)(4).
20. ☒ Other items or information: **Return Receipt Postcard; IDS includes ISR, IPER, Written Opinion; Cheque \$1040**



100903201071502  
10/088334

IC10 Rec'd PCT/PTO 14 MAR 2002  
Attorney Docket No.: 1473-071

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**  
**UNITED STATES RECEIVING OFFICE OF THE**  
**PATENT COOPERATION TREATY (35 USC §371)**

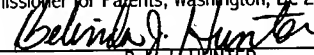
U.S. Applicant : VANDALI, Andrew E. *et al*  
U.S. Serial No. : To Be Assigned  
U.S. Filing Date : Herewith  
International Applicant : THE UNIVERSITY OF MELBOURNE  
International Application No. : PCT/AU00/01310  
International Filing Date : 25 October 2000 (25.10.2000)  
Earliest Claimed Priority : 26 October 1999 (26.10.1999)  
Title : EMPHASIS OF SHORT-DURATION TRANSIENT  
SPEECH FEATURES  
Examiner : To Be Assigned  
Group Art Unit : To Be Assigned

Box **PCT**  
Attn: National Phase Processing Division  
Assistant Commissioner Patent & Trademark  
Washington, DC 20231

**CERTIFICATE OF EXPRESS MAIL UNDER 37 CFR §1.10**

Express Mail Label # EL803248014US  
Date of Deposit March 14, 2002

I hereby certify that this paper and/or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service on the date indicated above and is address to: Box PCT, Attn: National Phase Processing Division, Assistant Commissioner for Patents, Washington, DC 20231.

  
Brenda HUNTER

**PRELIMINARY AMENDMENT**

S I R :

Prior to substantive examination, kindly amend the subject application as follows:

**IN THE CLAIMS:**

Please amend claims 3-5, 7, 8, and 10 as shown in the following section, "CLAIMS

IN CLEAN FORM."

## CLAIMS IN CLEAN FORM

3. The device of claim 1, wherein the detection of short-duration transitions in the rate of change profile causes up to about 14dB of gain to be applied to the sound signal, the amount of gain being selected according to the nature of the short-duration transition:

(i) a rapid increase followed by a decrease in the profile (a short-duration burst transition) cause a gain increase of up to about 14dB;

(ii) a rapid increase followed by a relatively constant level in the profile (an onset transition) causes a gain increase of up to about 6dB, and

(iii) a relatively constant level followed by a rapid decrease in the profile cause little or no increase in gain.

4. The device of claim 1, including a microphone, a pre-amplifier, a bank of N parallel filters, and means for applying a transient emphasis algorithm to the output of the filters.

5. The device of claim 1, wherein the signal in any filter channel denoted  $S_n(t)$ , where n denotes the filter channel number and t denotes time, is scaled according to:  $S'_n(t) = S_n(t) \times (1 + K_n \times G_n)$ , where  $G_n$  is the gain factor each channel and  $K_n$  is a gain modifier constant equal to about 2.

7. The device of claim 5, wherein the gain factor  $G_n$  is related to a function of the 2<sup>nd</sup> order derivative of the slow-varying envelope signal in each filter channel.

8. The device according to claim 5, wherein the gain factor for each filter channel is derived from the function:

$$G_n = (2 \times E_c - 2 \times E_p - E_r) / (E_c + E_p + E_r)$$

where  $E_c$ ,  $E_p$  and  $E_r$  are estimates of the current, past and future slow-varying envelope signal in each filter channel.

10. The device according to claim 8, wherein slow-varying envelope profiles exhibiting short-duration peaks of different peak levels but similar peak to valley ratios are amplified by similar amounts.

REMARKS

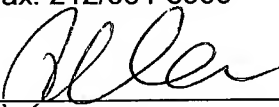
Amendments have been made hereinabove to the claims in order to eliminate multiple dependencies. Applicants state that no new matter is being added hereof.

Dated: 14 March 2002 (14.03.2002)  
New York, New York

Respectfully submitted,

GOTTLIEB RACKMAN & REISMAN PC  
Attorneys for Applicant  
270 Madison Avenue  
New York, New York 10016-0601  
Telephone: 212/684-3900  
Telefax: 212/684-3999

By: \_\_\_\_\_

  
Tiberiu WEISZ  
Reg. No. 29,876

## VERSION OF MARKINGS TO SHOW CHANGES MADE TO CLAIMS

3. The device of **[any preceding]** claim **1**, wherein the detection of short-duration transitions in the rate of change profile causes up to about 14dB of gain to be applied to the sound signal, the amount of gain being selected according to the nature of the short-duration transition:

(i) a rapid increase followed by a decrease in the profile (a short-duration burst transition) cause a gain increase of up to about 14dB;

(ii) a rapid increase followed by a relatively constant level in the profile (an onset transition) causes a gain increase of up to about 6dB, and

(iii) a relatively constant level followed by a rapid decrease in the profile cause little or no increase in gain.

4. The device of **[any preceding]** claim **1**, including a microphone, a pre-amplifier, a bank of N parallel filters, and means for applying a transient emphasis algorithm to the output of the filters.

5. The device of **[any preceding]** claim **1**, wherein the signal in any filter channel denoted  $S_n(t)$ , where n denotes the filter channel number and t denotes time, is scaled according to:  $S'_n(t) = S_n(t) \times (1 + K_n \times G_n)$ , where  $G_n$  is the gain factor each channel and  $K_n$  is a gain modifier constant equal to about 2.

7. The device of **[claims]** **claim 5 [or 6]**, wherein the gain factor  $G_n$  is related to a function of the 2<sup>nd</sup> order derivative of the slow-varying envelope signal in each filter channel.

8. The device according to **[claims] claim 5 [ , 6 or 7]**, wherein the gain factor for each filter channel is derived from the function:

$$G_n = (2 \times E_c - 2 \times E_p - E_r) / (E_c + E_p + E_r)$$

where  $E_c$ ,  $E_p$  and  $E_r$  are estimates of the current, past and future slow-varying envelope signal in each filter channel.

10. The device according to **[claims] claim 8 [and 9]**, wherein slow-varying envelope profiles exhibiting short-duration peaks of different peak levels but similar peak to valley ratios are amplified by similar amounts.



**EMPHASIS OF SHORT-DURATION TRANSIENT SPEECH**  
**FEATURES**

Field of the Invention

5           This invention relates to the processing of signals derived from sound stimuli, particularly for the generation of stimuli in auditory prostheses, such as cochlear implants and hearing aids, and in other systems requiring sound processing or encoding.

Background of the Invention

10           Various speech processing strategies have been developed for processing sound signals for use in stimulating auditory prostheses, such as cochlear prostheses and hearing aids. Such strategies focus on particular aspects of speech, such as formants. Other strategies rely on more general channelization and amplitude related selection, such as the Spectral Maxima Sound Processor  
15 (SMSP), strategy which is described in greater detail in Australian Patent No. 657959 by the present applicant, the contents of which are incorporated herein by cross reference.

          A recurring difficulty with all such sound processing systems is the provision of adequate information to the user to enable optimal perception of  
20 speech in the sound stimulus.

Summary of the Invention and Object

          It is an object of the present invention to provide a sound processing strategy to assist in perception of low-intensity short-duration speech features in the sound stimuli.

25           The invention provides a sound processing device having means for estimating the amplitude envelope of a sound signal in a plurality of spaced frequency channels, means for analyzing the estimated amplitude envelopes over time so as to detect short-duration amplitude transitions in said envelopes, means for increasing the relative amplitude of said short-duration amplitude  
30 transitions, including means for determining a rate of change profile over a predetermined time period of said short-duration amplitude transitions. and

- 2 -

means for determining from said rate of change profile the size of an increase in relative amplitude applied to said transitions in said sound signal to assist in perception of low-intensity short-duration speech features in said signal.

In a preferred form, the predetermined time period is about 60ms. The  
5 faster/greater the rate of change, on a logarithmic amplitude scale, of said short-duration amplitude transitions, the greater the increase in relative amplitude which is applied to said transitions. Furthermore, rate of change profiles corresponding to short-duration burst transitions receive a greater increase in relative amplitude than do profiles corresponding to onset transitions. In the  
10 present specification, a "burst transition" is understood to be a rapid increase followed by a rapid decrease in the amplitude envelope, while an "onset transition" is understood to be a rapid increase followed by a relatively constant level in the amplitude envelope.

The above defined Transient Emphasis strategy has been designed in  
15 particular to assist perception of low-intensity short-duration speech features for the severe-to-profound hearing impaired or Cochlear implantees. These speech features typically consist of: i) low-intensity short-duration noise bursts/frication energy that accompany plosive consonants; ii) rapid transitions in frequency of speech formants (in particular the 2nd formant, F2) such as  
20 those that accompany articulation of plosive, nasal and other consonants. Improved perception of these features has been found to aid perception of some consonants (namely plosives and nasals) as well as overall speech perception when presented in competing background noise.

The Transient Emphasis strategy is preferably applied as a front-end  
25 process to other speech processing systems, particularly but not exclusively, for stimulating implanted electrode arrays. The currently preferred embodiment of the invention is incorporated into the Spectral Maxima Sound Processor (SMSP) strategy, as referred to above. The combined strategy known as the Transient Emphasis Spectral Maxima (TESM) Sound Processor utilises the  
30 transient emphasis strategy to emphasise the SMSP's filter bank outputs prior to selection of the channels with the largest amplitudes.

- 3 -

As with most multi-channel speech processing systems, the input sound signal is divided up into a multitude of frequency channels by using a bank of band-pass filters. The signal envelope is then derived by rectifying and low-pass filtering the signal in these bands. Emphasis of short-duration transitions in the envelope signal for each channel is then carried out. This is done by: i) detection of short-duration (approximately 5 to 60 milliseconds) amplitude variations in the channel envelope typically corresponding to speech features such as noise bursts, formant transitions, and voice onset; and ii) increasing the signal gain during these periods. The gain applied is related to a function of the 2<sup>nd</sup> order derivative with respect to time of the slow-varying envelope signal (or some similar rule, as described below in the Description of Preferred Embodiment).

During periods of steady state or relatively slow varying levels in the envelope signal (over a period of approximately 60ms) no gain is applied. During periods where short-duration transition in the envelope signal are detected, the amount of gain applied can typically vary up to about 14dB. The gain varies depending of the nature of the short-duration transition which can be classified as either of the following. i) A rapid increase followed by a decrease in the signal envelope (over a period of no longer than approximately 60ms). This typically corresponding to speech features such as the noise-burst in plosive consonant or the rapid frequency shift of a formant in a consonant-to-vowel or vowel-to-consonant transition. ii) A rapid increase followed by relatively constant level in the signal envelope which typically corresponds to speech features such as the onset of voicing in a vowel. Short duration speech features classified according to i) are considered to be more important to perception than those classified according to ii) and thus receive relatively twice as much gain. Note, a relatively constant level followed by a rapid decrease in the signal envelope which corresponds to abruption of voicing/sound receive little to no gain.

30

Brief Description of the Drawings

In order that the invention may be more readily understood, one presently preferred embodiment of the invention will now be described with reference to the accompanying drawings in which:

5 Figure 1 is a schematic representation of the signal processing applied to the sound signal in accordance with the present invention, and

Figures 2 and 3 are comparative electrodiagrams of sound signals to show the effect of the invention.

Description of Preferred Embodiment

10 Referring to Figure 1, the presently preferred embodiment of the invention is described with reference to its use with the SMSP strategy. As with the SMSP strategy, electrical signals corresponding to sound signals received via a microphone 1 and pre-amplifier 2 are processed by a bank of N parallel filters 3 tuned to adjacent frequencies (typically  $N = 16$ ). Each filter channel  
15 includes a band-pass filter 4, then a rectifier 5 and low-pass filter 6 to provide an estimate of the signal amplitude (envelope) in each channel. In this embodiment a Fast Fourier Transform (FFT) implementation of the filter bank is employed. The outputs of the N-channel filter bank are modified by the transient emphasis algorithm 7 (as described below) prior to further processing  
20 in accordance with the SMSP strategy.

A running history, which spans a period of 60 ms. at 2.5 ms intervals, of the envelope signals in each channel, is maintained in a sliding buffer 8 denoted  $S_n(t)$  where the subscript n refers to the channel number and t refers to time relative to the current analysis interval. This buffer is divided up into three  
25 consecutive 20 ms time windows and an estimate of the slow-varying envelope signal in each window is obtained by averaging across the terms in the window. The averaging window provides approximate equivalence to a 2<sup>nd</sup>-order low-pass filter with a cut-off frequency of 45 Hz and is primarily used to smooth fine envelope structure, such as voicing frequency modulation, and unvoiced  
30 noise modulation. Averages from the three windows are therefore estimates of the past ( $E_p$ ) 9, current ( $E_c$ ) 10 and future ( $E_f$ ) 11 slow-varying envelope signal

- 5 -

with reference to the mid-point of the buffer  $S_n(t)$ . The amount of additional gain applied is derived from a function of the slow-varying envelope estimates as per Eq. (1). A derivation and analysis of this function can be found in Appendix A.

$$G = (2 \times E_c - 2 \times E_p - E_f) / (E_c + E_p + E_f) \quad (1)$$

The gain factor (G) 12 for each channel varies with the behaviour of the slow-varying envelope signals such that: (a) short-duration signals which consisted of a rapid rise followed by a rapid fall (over a time period of no longer than approximately 60 ms) in the slow-varying envelope signal produces the greatest values of G. For these types of signals, G could be expected to range from approximately 0 to 2. (b), The onset of long-duration signals which consist of a rapid rise followed by a relatively constant level in the envelope signal produces lower levels of G which typically range from 0 to 0.5. (c) A relatively steady-state or slow varying envelope signal produces negative value of G. (d) A relatively steady-state level followed by a rapid decrease in the envelope signal (i.e. cessation/offset of envelope energy) produces small (less than approximately 0.1) or negative values of G. Because negative values of G could arise, the result of Eq. (1) are limited at 13 such that it can never fall below zero as per Eq. (2).

$$\text{If } (G < 0) \text{ then } G = 0 \quad (2)$$

Another important property of Eq. (1) is that the gain factor is related to a function of relative differences, rather than absolute levels, in the magnitude of the slow-varying envelope signal. For instance, short-duration peaks in the slow-varying envelope signal of different peak levels but identical peak to valley ratios would be amplified by the same amount.

The gain factors for each channel ( $G_n$ ), where n denotes the channel number, are used to scale the original envelope signals  $S_n(t)$  according to Eq.(3), where  $t_m$  refers to the midpoint of the buffer  $S_n(t)$ .

$$S'_n(t_m) = S_n(t_m) \times (1 + K_n \times G_n) \quad (3)$$

A gain modifier constant ( $K_n$ ) is included at 14 for adjustment of the overall gain of the algorithm. In this embodiment,  $K_n = 2$  for all n. During

- 6 -

periods of little change in the envelope signal of any channel, the gain factor ( $G_n$ ) is equal to zero and thus  $S'_n(t_m) = S_n(t_m)$ , whereas, during periods of rapid-change,  $G_n$  could range from 0 to 2 and thus a total of 0 to 14 dB of gain could be applied. Note that because the gain is applied at the midpoint of the envelope signals, an overall delay of approximately 30 ms between the time from input to output of the transient emphasis algorithm is introduced. The modified envelope signals  $S'_n(t)$  at 15 replaces the original envelope signals  $S_n(t)$  derived from the filter bank and processing then continues as per the SMSP strategy. As with the SMSP strategy, M of the N channels of  $S'_n(t)$  having the largest amplitude at a given instance in time are selected at 16 (typically  $M = 6$ ). This occurs at regular time intervals and for the transient emphasis strategy is typically 2.5ms. The M selected channels are then used to generate M electrical stimuli 17 of stimulus intensity and electrode number corresponding to the amplitude and frequency of the M selected channels (as per the SMSP strategy). These M stimuli are transmitted to the Cochlear implant 19 via a radio-frequency link 18 and are used to activate M corresponding electrode sites.

Because the transient emphasis algorithm is applied prior to selection of spectral maxima, channels containing low-intensity short-duration signals, which: (a) normally fall below the mapped threshold level of the speech processing system; (b) or are not selected by the SMSP strategy due to the presence of channels containing higher amplitude steady-state signals: are given a greater chance of selection due to their amplification.

To illustrate the effect of the strategy on the coding of speech signals, stimulus output patterns, known as electrodograms (which are similar to spectrograms for acoustic signals), which plot stimulus intensity per channel as a function of time, were recorded for the SMSP and TESM strategies, and are shown in Figs. 2 & 3 respectively. The speech token presented in these recordings was /g o d/ and was spoken by a female speaker. The effect of the TESM strategy can be seen in the stimulus intensity and number of electrodes representing the noise burst energy in the initial stop /g/ (point A). The onset of the formant energy in the vowel /o/ has also been emphasised slightly (point B).

- 7 -

Most importantly, stimuli representing the second formant transition from the vowel /o/ to the final stop /d/ are also higher in intensity (point C), as are those coding the noise burst energy in the final stop /d/ (point D).

## 5 APPENDIX A: TESM Gain Factor

To derive a function for the gain factor (G) 12 for each channel in terms of the slow-varying envelope signal the following criteria were used. Firstly, the gain factor should be related to a function of the 2<sup>nd</sup> order derivative of the  
10 slow-varying envelope signal. The 2<sup>nd</sup> order derivative is maximally negative for peaks (and maximally positive for valleys) in the slow-varying envelope signal and thus it should be negated; Eq. (A1).

$$G \propto 2 \times E_c - E_p - E_f \quad (A1)$$

Secondly, for the case when the 'backward' gradient (i.e.  $E_c - E_p$ ) is  
15 positive but small, significant gain as per Eq. (A1) can result when  $E_f$  is small (i.e. at the cessation (offset) of envelope energy for a long-duration signal). This effect is not desirable and can be minimised by reducing the backward gradient to near zero or less (i.e. negative) in cases when it is small. However, when the backward gradient is large, Eq. (A1) should hold. A simple solution is  
20 to scale  $E_p$  by 2. A function for the 'modified' 2<sup>nd</sup> order derivative is given in Eq. (A2). As  $E_p$  approaches  $E_c$ , G approaches  $-E_f$  rather than  $E_c - E_f$ , as in Eq. (A1) and thus the gain factor approaches a small or negative value. However for  $E_p \ll E_c$ , G approaches  $2 \times E_c - E_f$ , which is identical to the limiting condition for Eq. (A1).

$$25 \quad G \propto 2 \times E_c - 2 \times E_p - E_f \quad (A2)$$

Thirdly, because we are interested in providing gain based on relative rather than absolute differences in the slow-varying envelope signal, the gain factor should be normalised with respect to the average level of slow-varying envelope signal as per Eq. (A3). The effect of the numerator in Eq. (A3)  
30 compresses the linear gain factor as defined in Eq. (A2) into a range of 0 to 2. The gain factor is now proportional to the modified 2<sup>nd</sup> order derivative and

- 8 -

inversely proportional to the average level of the slow-varying envelope channel signal.

$$G = (2 \times E_c - 2 \times E_p - E_f) / (E_c + E_p + E_f) \quad (A3)$$

Finally, the gain factor according to Eq. (A3) can fall below zero when  $E_c < E_p + E_f/2$ . Thus, Eq. (A4) is imposed on  $G_n$  so that the gain is always greater than or equal to zero.

$$\text{If } (G < 0) \text{ then } G = 0 \quad (A4)$$

An analysis of the limiting cases for the gain factor can be used to describe its behaviour as a function of the slow-varying envelope signal. For the limiting case when  $E_p$  is much smaller than  $E_c$  (i.e. during a period of rapid-rise in the envelope signal), Eq. (A3) reduces to:

$$G = (2 \times E_c - E_f) / (E_c + E_f) \quad (A5)$$

In this case, if  $E_f$  is greater than  $E_c$  and approaches  $2 \times E_c$  (i.e. during a period of steady rise in the slow-varying envelope signal),  $G$  approaches zero. If  $E_f$  is similar to  $E_c$  (i.e. at the end a period of rise for a long-duration signal),  $G$  is approximately 0.5. If  $E_f$  is a lot smaller than  $E_c$  (i.e. at the apex of a rapid-rise which is immediately followed by a rapid fall as is the case for short-duration peak in the envelope signal),  $G$  approaches 2, which is the maximum value possible for  $G$ .

For the limiting case when  $E_f$  is much smaller than  $E_c$ , Eq. (A3) reduces to:

$$G = (2 \times E_c - 2 \times E_p) / (E_c + E_p) \quad (A6)$$

In this case, if  $E_c$  is similar to  $E_p$  (i.e. cessation/offset of envelope for a long-duration signal),  $G$  approaches zero. If  $E_c$  is much greater than  $E_p$  (i.e. at a peak in the envelope),  $G$  approaches the maximum gain of 2.

When dealing with speech signals, intensity is typically defined to on a log (dB) scale. It is thus convenient to view the applied gain factor in relation to the gradient of the log-magnitude of the slow-varying envelope signal. Eq. (A3) can be expressed in terms of ratios of the slow-varying envelope signal estimates. Defining the backward magnitude ratio as  $R_b = E_c/E_p$  and the forward magnitude ratio  $R_f = E_f/E_c$  gives Eq. (A7).



- 9 -

$$G = (2 \times R_b - 2 - R_b \times R_f) / (R_b + 1 + R_b \times R_f) \quad (A7)$$

The forward and backward magnitude ratios are equivalent to log-magnitude gradients and can be as defined as the difference between log-magnitude terms, i.e.  $F_g = \log(E_f) - \log(E_c)$  and  $B_g = \log(E_c) - \log(E_p)$  respectively. The relationship between gain factor and forward and backward log-magnitude gradients is shown in Fig. A1. Linear gain is plotted on the ordinate and backward log-magnitude gradient (in dB) is plotted on the abscissa. The gain factor is plotted for different levels of the forward log-magnitude gradient in each of the curves. For any value of the forward log-magnitude gradient, the gain factor reaches some maximum when the backward log-magnitude gradient is approximately 40 dB. The maximum level is dependent on the level of the forward log-magnitude gradient. For the case where the forward log-magnitude gradient is 0 dB, as shown by the dotted line (i.e. at the end a period of rise for a long-duration signal where  $E_f = E_c$ ) the maximum gain possible is 0.5. For the limiting case where the forward log-magnitude gradient is infinitely steep as shown by the dashed line (i.e. rapid-fall in envelope signal where  $E_f \ll E_c$ ), the maximum gain possible is 2.0. The limiting case for the forward log-magnitude gradient is reached when its gradient is approximately - 40 dB.

- 10 -

## CLAIMS:

1. A sound processing device having means for estimating the amplitude envelope of a sound signal in a plurality of spaced frequency channels, means  
5 for analyzing the estimated amplitude envelopes over time so as to detect short-duration amplitude transitions in said envelopes, means for increasing the relative amplitude of said short-duration amplitude transitions, including means for determining a rate of change profile over a predetermined time period of said short-duration amplitude transitions, and means for determining from said  
10 rate of change profile the size of an increase in relative amplitude applied to said transitions in said sound signal to assist in perception of low-intensity short-duration speech features in said signal.
2. The device of claim 1, wherein the predetermined time provided is  
15 approximately 60ms, and the faster/greater the rate of change, on a logarithmic amplitude scale, of said short-duration amplitude transitions, the greater the increase in relative amplitude which is applied to said transitions.
3. The device of any preceding claim, wherein the detection of short-  
20 duration transitions in the rate of change profile causes up to about 14dB of gain to be applied to the sound signal, the amount of gain being selected according to the nature of the short-duration transition:
  - (i) a rapid increase followed by a decrease in the profile (a short-  
duration burst transition) cause a gain increase of up to about 14dB;
  - 25 (ii) a rapid increase followed by a relatively constant level in the profile (an onset transition) causes a gain increase of up to about 6dB, and
  - (iii) a relatively constant level followed by a rapid decrease in the profile causes little or no increase in gain.

- 11 -

4. The device of any preceding claim, including a microphone, a pre-amplifier, a bank of N parallel filters, and means for applying a transient emphasis algorithm to the output of the filters.

5. The device of any preceding claim, wherein the signal in any filter channel denoted  $S_n(t)$ , where n denotes the filter channel number and t denotes time, is scaled according to:  $S'_n(t) = S_n(t) \times (1 + K_n \times G_n)$ , where  $G_n$  is the gain factor for each channel and  $K_n$  is a gain modifier constant equal to about 2.

6. The device of claim 5, wherein a buffer maintains a history of the envelope signal  $S_n(t)$  in each filter channel for which an estimate of the slow-varying envelope signal is derived by an averaging window which provides an appropriate equivalence to a 2<sup>nd</sup>-order low-pass filter with a cut off frequency of about 45Hz causing smoothing of the fine profile structure.

7. The device of claims 5 or 6, wherein the gain factor  $G_n$  is related to a function of the 2<sup>nd</sup>-order derivative of the slow-varying envelope signal in each filter channel.

8. The device according to claims 5, 6 or 7, wherein the gain factor for each filter channel is derived from the function:

$$G_n = (2 \times E_c - 2 \times E_p - E_f) / (E_c + E_p + E_f)$$

where  $E_c$ ,  $E_p$  and  $E_f$  are estimates of the current, past and future slow-varying envelope signal in each filter channel.

9. The device of claim 8, wherein the additional factor gain  $G_n$  applied to the sound signal can range from about 0 to 2 for a slow-varying envelope profile having a rapid rise followed by a rapid fall, about 0 to 0.5 for a profile having a rapid rise followed by a relatively constant level, less than 0.1 for a steady state level followed by a rapid decrease in the profile, and about 0 for a steady state level or slowly varying profile.

10. The device according to claims 8 and 9, wherein slow-varying envelope profiles exhibiting short-duration peaks of different peak levels but similar peak to valley ratios are amplified by similar amounts.

5

## (12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
3 May 2001 (03.05.2001)

PCT

(10) International Publication Number  
**WO 01/31632 A1**

(51) International Patent Classification<sup>7</sup>: **G10L 15/00 //**  
101:02, H04R 25/00

(72) Inventors; and  
(75) Inventors/Applicants (for US only): **VANDALI, Andrew, E** [AU/AU]; 14 Arncliffe Boulevard, Greenvale, VIC 3059 (AU). **CLARK, Graeme, M.** [AU/AU]; 13 Banoon Road, Eltham, VIC 3095 (AU).

(21) International Application Number: **PCT/AU00/01310**

(22) International Filing Date: 25 October 2000 (25.10.2000)

(74) Agent: **FREEHILLS CARTER SMITH BEADLE**; 101 Collins Street, Melbourne, VIC 3000 (AU).

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
PQ 3667 26 October 1999 (26.10.1999) AU

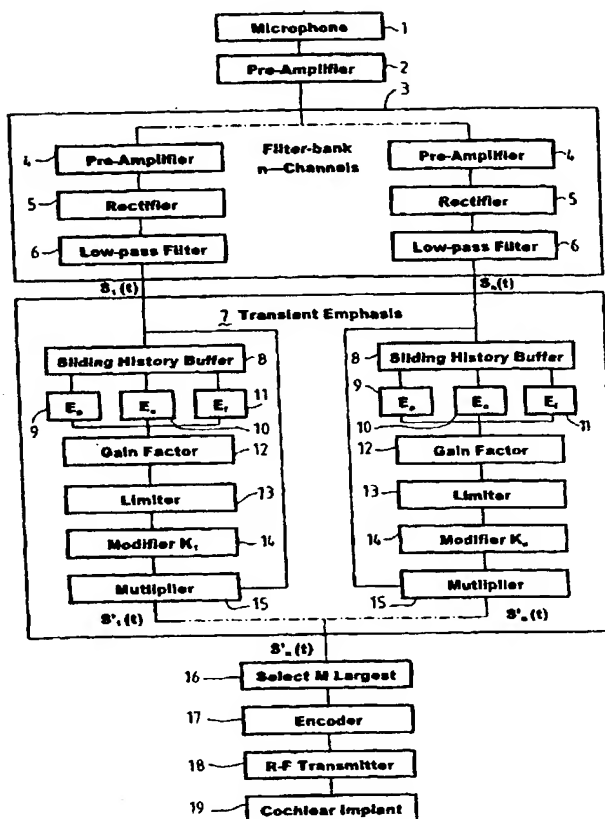
(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(71) Applicant (for all designated States except US): **THE UNIVERSITY OF MELBOURNE** [AU/AU]; Grattan Street, Parkville, VIC 3052 (AU).

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian

[Continued on next page]

(54) Title: **EMPHASIS OF SHORT-DURATION TRANSIENT SPEECH FEATURES**



(57) Abstract: A sound processor including a microphone (1), a pre-amplifier (2), a bank of N parallel filters (3), means for detecting short-duration transitions in the envelope signal of each filter channel, and means for applying gain to the outputs of these filter channels in which the gain is related to a function of the second-order derivative of the slow-varying envelope signal in each filter channel, to assist in perception of low-intensity short-duration speech features in said signal.

WO 01/31632 A1

7

WO 01/31632 A1

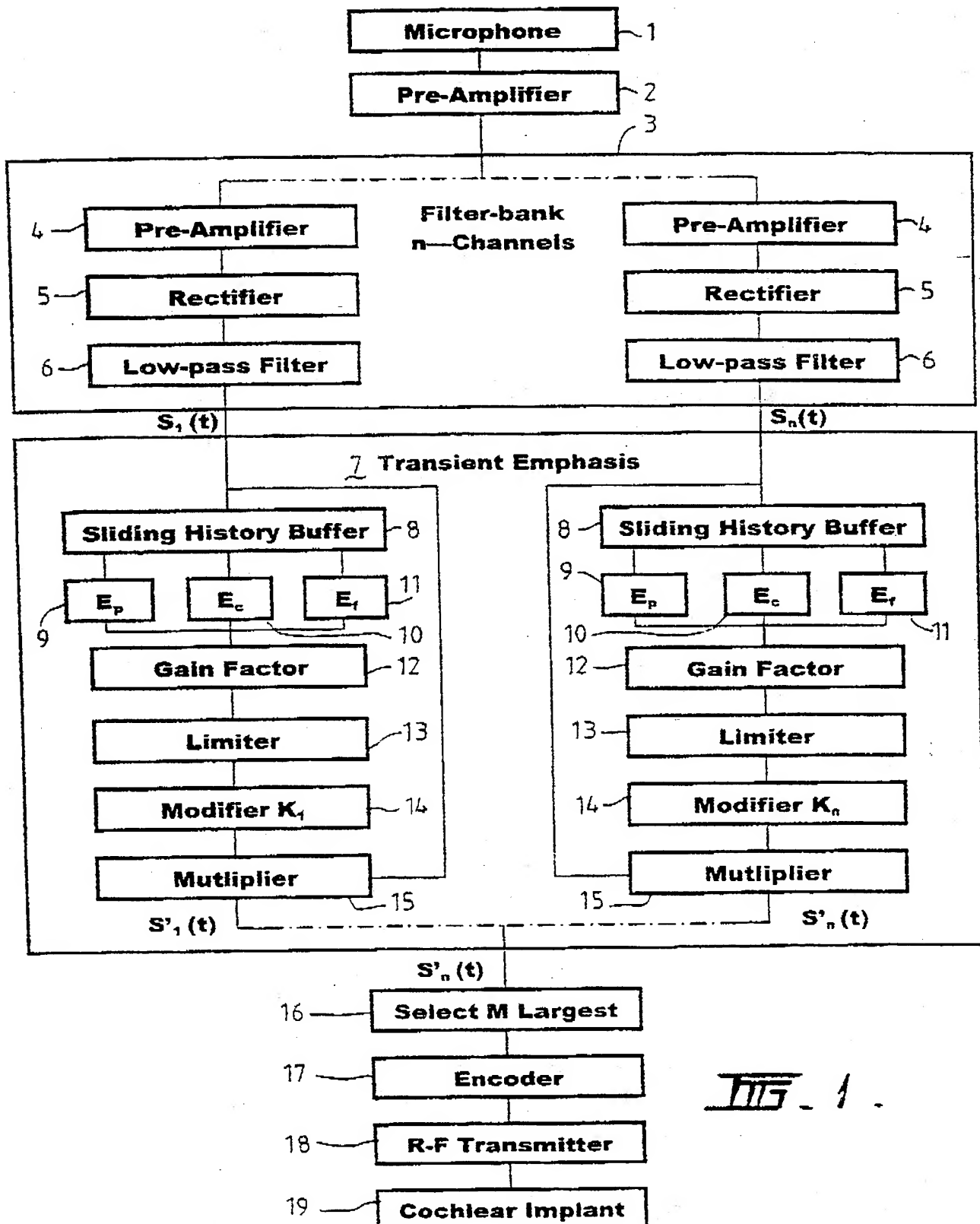


patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European  
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,  
IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG,  
CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

**Published:**

With international search report.

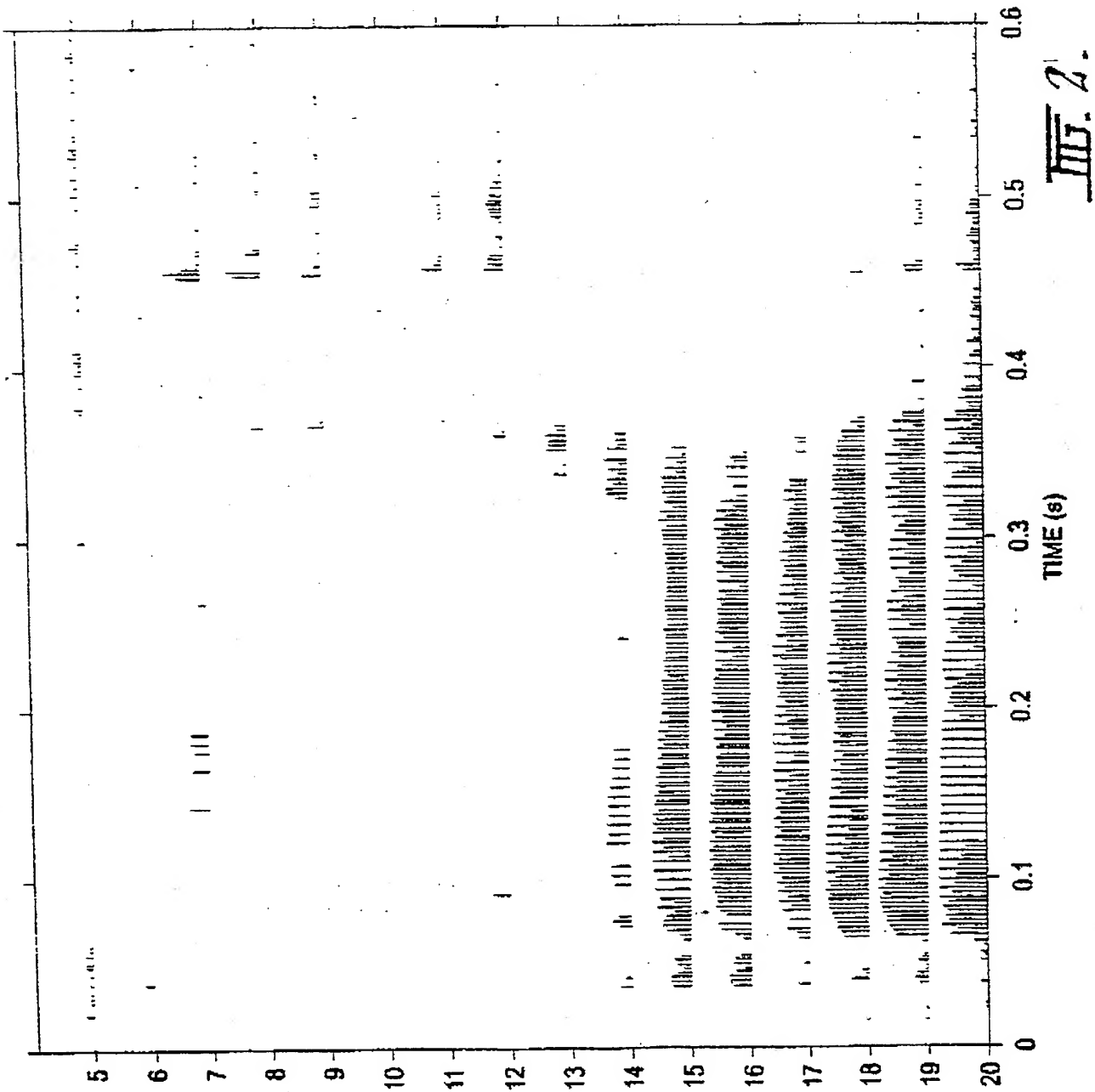


10/088334

WO 01/31632

PCT/AU00/01310

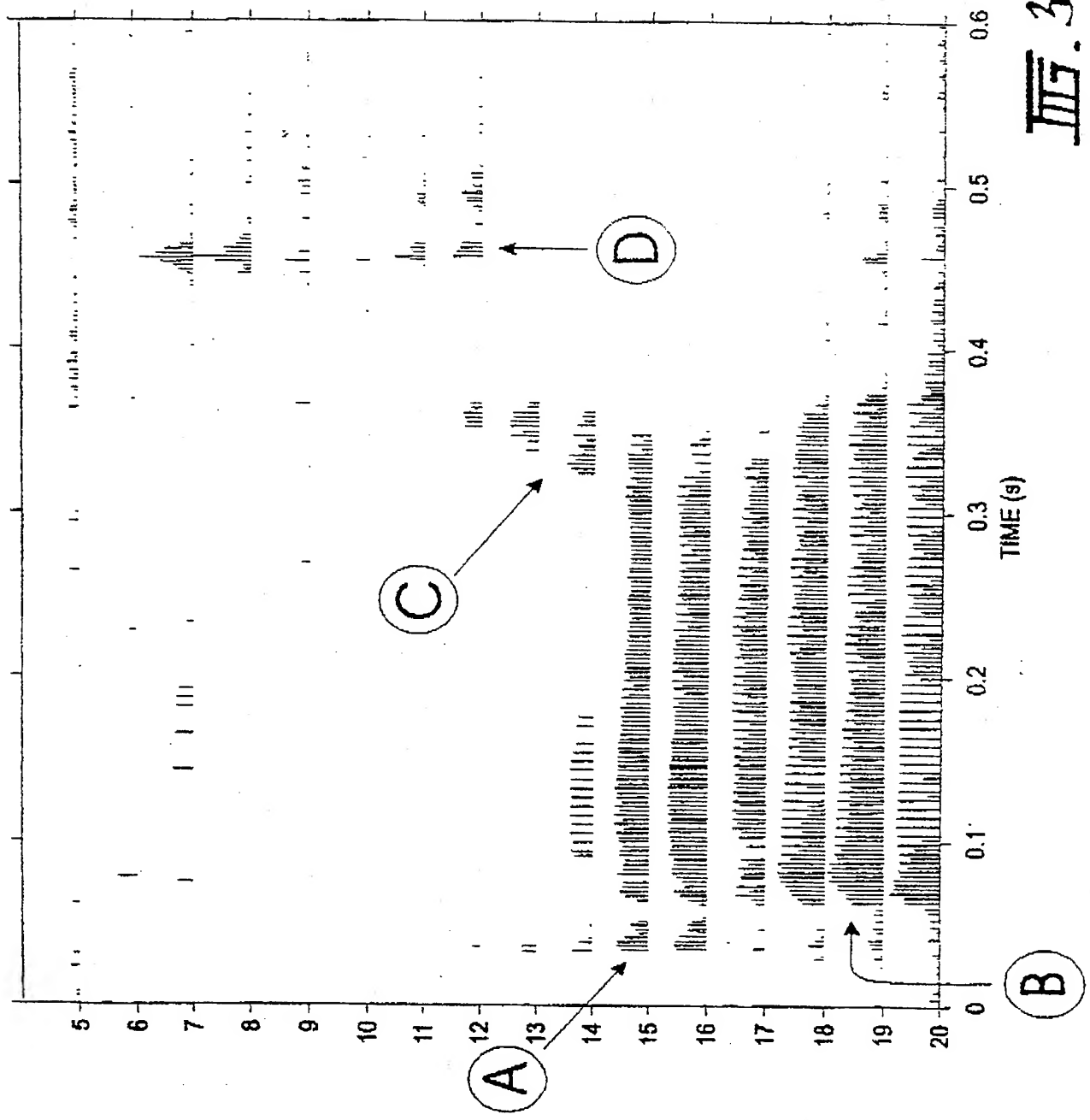
2 / 4



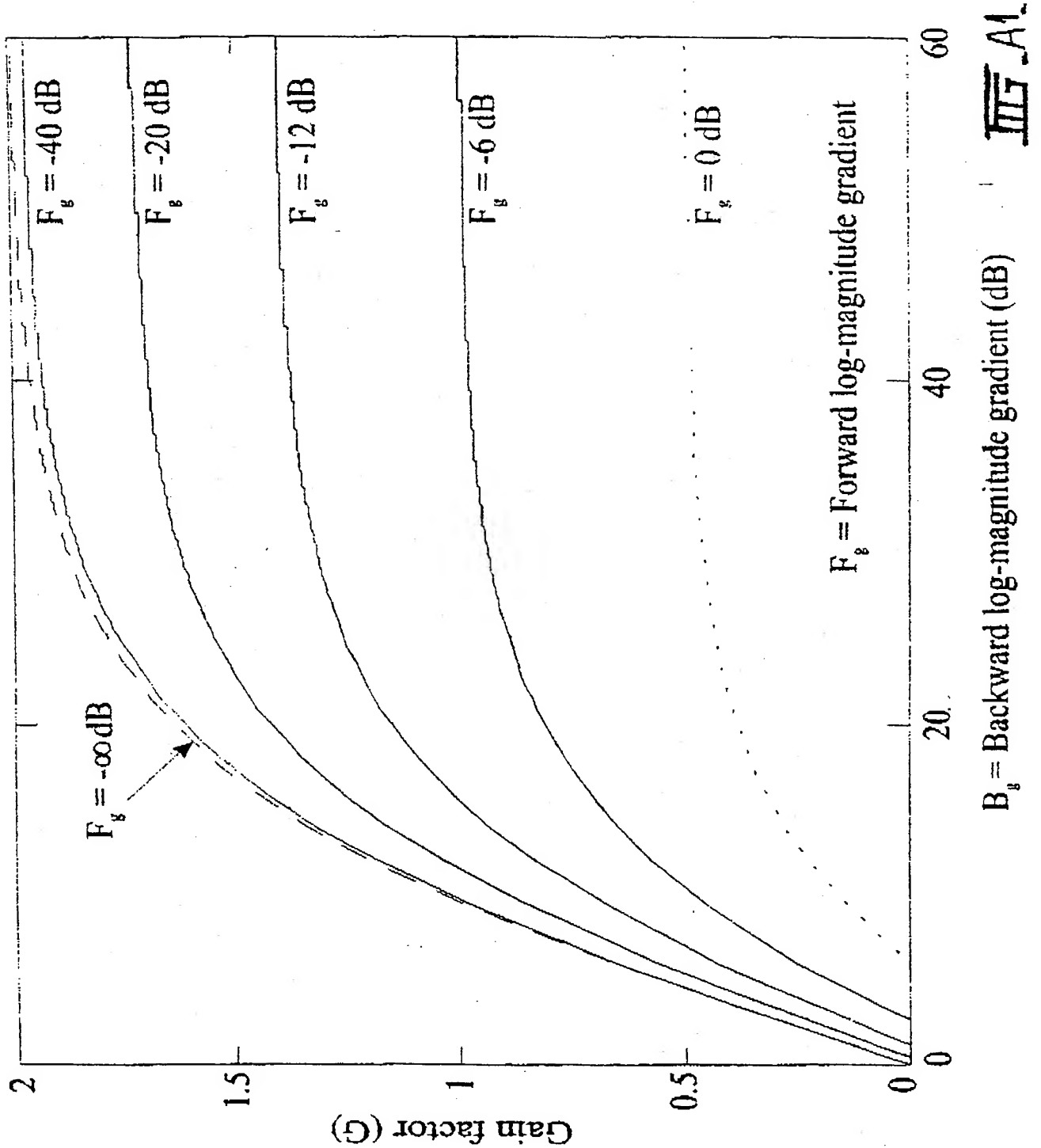


WO 01/31632

PCT/AU00/01310



III.3



1473-071

## Combined Declaration and Power of Attorney

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled **EMPHASIS OF SHORT-DURATION TRANSIENT SPEECH FEATURES** the specification of which

(check one) ☐ is attached hereto. ☒ was filed on March 14, 2002 as United States Application Serial No. 10/088,334 or PCT International Application No. \_\_\_\_\_ and was amended on \_\_\_\_\_ (if applicable).

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations, §1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, §119 (a)-(d) or 365(b) of any foreign application(s) for patent or inventor's certificate, or 365(a) of any PCT international application which designated at least one country other than the United States of America, or Title 35, United States Code, §371, listed below and have also identified below, by checking the appropriate box, any foreign application for patent or inventor's certificate, or of any PCT application having a filing date before that of the application on which priority is claimed:

Prior Foreign Application(s)			Priority claimed	
<u>PCT/AU00/01310</u> ✓	<u>Australia</u> ✓	<u>25 October 2000</u> ✓	<input checked="" type="checkbox"/>	<input type="checkbox"/>
(Number)	(Country)	Day/month/year filed	Yes	No
<u>PQ3667</u> ✓	<u>Australia</u> ✓	<u>26 October 1999</u> ✓	<input checked="" type="checkbox"/>	<input type="checkbox"/>
(Number)	(Country)	Day/month/year filed	Yes	No

I hereby claim the benefit under Title 35, United States Code, § 119(e) of any United States provisional application(s) listed below:

\_\_\_\_\_  
(Application No.) (Filing Date)

\_\_\_\_\_  
(Application No.) (Filing Date)

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations, §1.56 which occurred between the filing date of the prior application and the national or PCT international filing date of this application:

\_\_\_\_\_  
(Application No.) (Filing date) (Status - patented, pending, abandoned)

\_\_\_\_\_  
(Application No.) (Filing date) (Status - patented, pending, abandoned)

And I hereby appoint

George Gottlieb (Reg.No. 22,035)	Jeffrey M. Kaden (Reg.No. 31,268)
Michael I. Rackman (Reg.No. 20,639)	Amy B. Goldsmith (Reg.No. 33,700)
James Reisman (Reg.No. 22,007)	Norbert P. Holler (Reg.No. 17,816)
Barry A. Cooper (Reg.No. 25,204)	Tiberlu Weisz (Reg.No. 29,876)
David S. Kashman (Reg.No. 28,725)	Marla A. Savio (Reg.No. 31,565)
Allen I. Rubenstein (Reg.No. 27,673)	Raymond B. Churchill (Reg.No. 44,617)

whose address is Gottlieb, Rackman & Reisman, P.C., 270 Madison Avenue, New York NY, 10016-0604 (telephone (212) 684-3900), jointly and severally, as my attorneys and/or agents, with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith.

Direct all correspondence and telephone calls to: Tiberlu WEISZ, Esq. at the address and telephone number shown above.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

1-00 Full name of sole or first inventor VANDALI, Andrew E. Date 25 March 2002  
 Inventor's Signature \_\_\_\_\_  
 Residence 14 Arncliffe Boulevard, Greenvale, Victoria 3059, Australia  
 Citizenship Commonwealth of Australia  
 Post Office Address Greenvale, Victoria 3059, Australia

2-02 Full name of second joint inventor, if any CLARK, Graeme M. Date 1/5/02  
 Inventor's Signature \_\_\_\_\_  
 Residence 13 Banook Road, Eltham, Victoria 3095, Australia  
 Citizenship Commonwealth of Australia  
 Post Office Address Eltham, Victoria 3095, Australia

Full name of third joint inventor, if any \_\_\_\_\_ Date \_\_\_\_\_  
 Inventor's Signature \_\_\_\_\_  
 Residence \_\_\_\_\_  
 Citizenship \_\_\_\_\_  
 Post Office Address \_\_\_\_\_

Full name of fourth joint inventor, if any \_\_\_\_\_ Date \_\_\_\_\_  
 Inventor's Signature \_\_\_\_\_  
 Residence \_\_\_\_\_  
 Citizenship \_\_\_\_\_  
 Post Office Address \_\_\_\_\_

LF1A.12/99